Talanta 267 (2024) 125265

Contents lists available at ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

Model inversion and three-way decompositions in the analytical quality by design strategy for the determination of phthalates BY HS-SPME-GC-MS

Lucía Valverde-Som [a], Ana Herrero [a], Celia Reguera [a], Luis Antonio Sarabia [b], María Cruz Ortiz [a,*], María Sagrario Sánchez [b]

[a] Department of Chemistry, Faculty of Sciences, Universidad de Burgos, Plaza Misael Bañuelos s/n, 09001 Burgos Spain
[b] Department of Mathematics and Computation, Faculty of Sciences, Universidad de Burgos, Plaza Misael Bañuelos s/n, 09001 Burgos Spain

ARTICLE INFO

Handling editor: Kin-ichi Tsunoda

Keywords:
Bottled natural and flavoured mineral water
Phthalates
HS-SPME-GC-MS
PLS2 model inversion
PARAFAC2
Analytical Quality by Design

ABSTRACT

In this work, strategies within Analytical Quality by Design (AQbD) with tools of the Process Analytical Technology (PAT) were used in the development of a head space-solid phase microextraction-gas chromatography-mass spectrometry (HS-SPME-GC-MS) procedure for the multiresidue analysis of four phthalic acid esters, benzyl butyl phthalate, bis(2-ethylhexyl) phthalate, dibutyl phthalate and diethyl phthalate (regulated by Commission Regulation (EU) No 10/2011). The approach is based on the fact that the intended quality of the resulting chromatograms is defined in terms of the loadings on the sample mode of a Parallel Factor Analysis (PARAFAC2) decomposition. These loadings are the ones used for the inversion of a Partial Least Squares (PLS2) prediction model that has been previously fitted. The inversion gives the experimental conditions that represent a compromise solution in terms of the desired or target values of the responses (Critical Quality Attributes, CQAs), while guaranteeing that these experimental conditions are inside the experimental domain of the Control Method Parameters (CMPs).

This strategy results in experimental conditions of extraction time and temperature that lead to a chromatogram of predefined quality for the four analytes together, with the subsequent saving of time and energy. The experimental conditions achieved have been experimentally verified and figures of merit of the analytical method have been determined. The method has been applied to a case study, bottled natural and flavoured mineral water. Concentrations around 0.3 μg L^{-1} of dibutyl phthalate have been found in 5 of the 22 bottles of water analysed.

1. Introduction

The term Analytical Quality by Design (AQbD) refers to the Quality by Design (QbD) paradigm when it is applied to the development of analytical methods or procedures. This is so because analytical methods can be seen as real "processes" with their own inputs (usually referred to as Control Method Parameters -CMPs) that influence the outputs. The quality of these outputs is evaluated in terms of the customarily named Critical Quality Attributes (CQAs). The idea is that the quality of the resulting "product" (the outputs) should be built into the process.

In practice, when planning the development of an analytical method, the focus is indeed on the intended quality, that is, the characteristics of the CQAs that the method should meet. In AQbD, CQAs are also called Critical Method Attributes (CMAs) or Critical Analytical Attributes (CAAs) [1–3]. In any case, their intended quality is precisely defined in the Analytical Target Profile (ATP). The problem then becomes finding the values of the CMPs so that the quality of the resulting output matches the desired quality in the CQAs.

To tackle the problem, a prediction model relating CMPs and CQAs is needed. If $\mathbf{x}$ denotes a vector of settings of some CMPs and f is the prediction model, the common use of f is to compute the vector of predicted values for the CQAs, $\hat{\mathbf{y}} = f(\mathbf{x})$. Now, if $\mathbf{y}_d$ is the vector with the intended CQAs, this means that it is necessary to find or construct a vector of CMPs, $\hat{\mathbf{x}}$, such that $\hat{\mathbf{y}} = f(\hat{\mathbf{x}}) = \mathbf{y}_d$. In other words, the inversion of f for $\mathbf{y}_d$.

Although most of the approaches in the literature use polynomial models (usually quadratic models fitted with response surface designs), the proposal in the present work includes the use of latent variable prediction models to take into account the expected high correlation among responses and CMPs, precisely, Partial Least Squares (PLS or PLS2) models.

* Corresponding author.
E-mail address: mcortiz@ubu.es (M.C. Ortiz).

https://doi.org/10.1016/j.talanta.2023.125265
Received 24 July 2023; Received in revised form 29 September 2023; Accepted 1 October 2023
Available online 3 October 2023

Abbreviations

(AQbD) Analytical Quality by Design
(ATP) analytical target profile
(BBP) benzyl butyl phthalate
(DEHP) bis(2-ethylhexyl) phthalate
(CCβ) capability of detection
(CMPs) control method parameters
(CORCONDIA) core consistency diagnostic
(CMAs) critical method attributes
(CQAs) critical quality attributes
(CCα) decision limit
(DBP) dibutyl phthalate
(DEP) diethyl phthalate
(DiBP-d_4) diisobutyl phthalate-3, 4, 5, 6-d_4
(HS-SPME-GC-MS) head space-solid phase microextraction-gas chromatography-mass spectrometry
(IS) internal standard
(PARAFAC) parallel factor analysis
(PLS) partial least squares
(PAEs) phthalatic acid esters
(RPET) recycled polyethylene terephthalate
(RMSECV) root mean squared error in cross validation

In fact, when using PLS2 in the usual multi-response situation, there is not a single setting of experimental conditions that serve for all responses. For that reason, unlike the usual procedures, the inversion is made by looking for the Pareto optimal front that accounts for the trade-off among the CQA, instead of optimizing each response separately.

This approach has already been successfully applied to obtain a preset ATP in the determination of five bisphenols by HPLC-DAD [4] and eight triazines in surface waters by means of SPE-HPLC-DAD [5]. In both cases, the desired characteristics of the intended chromatogram are defined through the proper resolution for consecutive peaks and short time of analysis, whereas the CMPs are conditions of the composition and flow rate of a ternary mobile phase. Also, in Ref. [6] the PLS2 model inversion has been applied to obtain the method operable design region (MODR) in the determination of ten polycyclic aromatic hydrocarbons by HPLC-FLD.

The interest of the AQbD (or of QbD) has grown in recent years in fields not only of pharmacology, toxicology, pharmacy, but also in the field of chemistry. The supplementary material of Ref. [7] contains a recent summary of several of these publications, all of them related to chromatography, mostly are liquid chromatography with different detectors (DAD, FLD). Also, the review in Ref. [8] devotes its Table 2 for HPLC, in all the cases using experimental designs to fit and optimize with OLS (ordinary least squares) linear or quadratic models. Another recent review [9] focuses on the new trends with chiral capillary electrophoresis.

However, a search in SCOPUS for papers that apply AQbD with GC-MS gave only four of them, Refs. [10–13], two of which [10,13] use SPME-GC-MS but, again, all of them with polynomial models based on experimental designs (central composite in Refs. [10–12] and fractional factorial in Ref. [13]). The CMPs vary between 2 and 4, while the CQAs vary between 2 and 16. In all cases, the peak areas (or concentrations), the resolutions between peaks [11,12] or the intermediate precision [13] are included. Details about the analytes, the analytical technique, the experimental design used, and other characteristics are summarized in Table S1 of the supplementary material.

In any case, no one of them uses PARAFAC, which is the proposal here in the AQbD context, and with the focus on maintaining the CQAs. Precisely, the present work proposes the inversion of a PLS prediction model through the sample loadings of a parallel factor analysis (PARAFAC2) decomposition (linearly related to the concentration of the analytes) to find adequate experimental conditions in the development of a head space-solid phase microextraction-gas chromatography-mass spectrometry (HS-SPME-GC-MS) procedure for the determination of four phthalic acid esters (PAEs), benzyl butyl phthalate (BBP), bis(2-ethylhexyl) phthalate (DEHP), dibutyl phthalate (DBP) and diethyl phthalate (DEP), in bottled natural and flavoured mineral water.

These PAEs are commonly used plasticisers in polyethylene and other plastic materials employed as additives in toys, plastic cosmetics, medical devices, or food packaging, among others [14,15]. Numerous examples in the literature [16,17] refer to their presence in bottled waters due to migration processes during storage, although it has been also suggested that they can be biosynthesised in nature [18]. Their presence in these waters may pose a health risk [19,20], in fact, BBP, DEHP and DBP are listed in Annex XIV to Regulation (EC) No 1907/2006 [21] as substances toxic for reproduction, category 1B. The use of PAEs in materials and articles intended to come into contact with food is regulated by Commission Regulation (EU) No 10/2011 [22].

Chromatographic techniques are the most widely used techniques for the identification and quantification of PAEs [23,24]. Because these compounds are often found at trace levels, sample pre-treatment or extraction methods are usually necessary prior to instrumental analysis. Among the latter, solid-phase microextraction (SPME) is considered as a green analytical chemistry technique [25] and has already shown its usefulness in the determination of phthalates in aqueous samples [26], and also other types of complex matrices, such as milk [27], vegetable oil [28] or wine [29], in all the cases coupled to GC-MS.

When multi-way decomposition techniques are used with this type of chromatographic data, a very powerful analytical tool is available, which is particularly useful in regulated frameworks [30,31]. Techniques such as PARAFAC [32], with the second order advantage, allow determinations to be made even in the presence of unexpected and unknown interferents, making the unequivocal identification of the target compounds possible [26,30,31], as required by the regulations [33].

The novelty of the approach developed in this work is that the intended quality of the resulting chromatograms, i.e. the CQAs, is defined in terms of the PARAFAC (PARAFAC2) loadings on the sample mode, which are the ones to be used for the inversion of the prediction model that must be previously fitted.

2. Material and methods

2.1. Chemicals

Benzyl butyl phthalate (BBP, CAS no. 85-68-7, analytical standard, ≥ 98 % purity), bis(2-ethylhexyl) phthalate (DEHP, CAS no. 117-81-7, analytical standard, ≥ 98 % purity), dibutyl phthalate (DBP, CAS no. 84-74-2, analytical standard, ≥ 98 % purity), diethyl phthalate (DEP, CAS no. 84-66-2, 99.5 % purity), and diisobutyl phthalate-3,4,5,6-d_4 (DiBP-d_4, CAS no. 358730-88-8, ≥ 98 % purity), used as internal standard (IS), were purchased from Sigma-Aldrich (Steinheim, Germany). Granular sodium chloride (NaCl, CAS no. 7647-14-5), in glass container, was purchased from Avantor Performance Materials (Center Valley, PA). A Milli-Q gradient A10 water purification system from Millipore (Bedford, MA, USA) was used to obtain Milli-Q water.

Acetone (CAS no. 67-64-1), methanol (CAS no. 67-56-1) and n-hexane (CAS no. 110-54-3), for liquid chromatography Lichrosolv® were from Merck KGaA (Darmstadt, Germany), which were used to clean the glass materials.

Helium (99.999 % purity, ALPHAGAZ™ 1, Air Liquide, Madrid, Spain) was used as the carrier gas.

2.2. Standard solutions

Stock solutions of DEP and DBP at 2300 mg L^{-1}, of BBP and DEHP at

2000 mg L^{-1}, and of DiBP-d_4 at 500 mg L^{-1} were prepared individually in methanol. All intermediate standard solutions were also prepared in methanol, so that the final aqueous standards contained 0.5 % methanol. Final standards, with 10 % (w/w) sodium chloride and the required concentrations of the target compounds and IS, were prepared from the intermediate solutions.

All stock solutions were gravimetrically controlled, to verify that the solvent had not evaporated, and stored in crimp vials, protected from light, at 4 °C. All intermediate and final solutions were prepared daily.

2.3. Bottled natural and flavoured mineral water samples

Thirteen bottled flavoured mineral waters (five of them carbonated waters) and nine bottled natural mineral waters (two of them carbonated waters) were analysed (W1–W22), 22 samples in total. Commercial mineral waters of different brands, the characteristics of which are listed in Table 1, were purchased at local stores (Burgos, Spain).

Water samples were prepared so that they contained 10 % sodium chloride and 0.4 μg L^{-1} of IS (as well as 0.5 % methanol). Carbonated water samples were previously placed in an ultrasound bath, until they are completely degassed.

2.4. Instrumentation

Analyses were performed on an Agilent 6890 N gas chromatograph with a split-splitless injector, coupled to an Agilent 5975 mass spectrometer detector with a single quadrupole mass analyser (Agilent Technologies, Palo Alto, CA, USA). A glass liner for SPME (0.75 mm ID) was used. Chromatographic separation was achieved with a capillary column with dimensions of 30 m × 0.25 mm inner diameter × 0.25 μm film thickness and coated with an Agilent HP-5MS Ultra Inert column ((5 %-phenyl)-methylpolysiloxane stationary phase, J&W Scientific, Folsom, CA, USA). A TriPlus autosampler equipped with a SPME module (Thermo Scientific, Milan, Italy) was used by the SPME procedure. A divinylbenzene/carboxen on polydimethylsiloxane (DVB/CAR/PDMS, 50/30 μm film thickness) was used as the extraction fibre which was supplied by Supelco (Bellefonte, PA, USA).

Table 2
Variance explained in predictor **X** and response **Y** variables when adding latent variables (LV) in the PLS2 model.

LV number	Variance explained in X (%)	Cumulative variance explained in X (%)	Variance explained in Y (%)	Cumulative variance explained in Y (%)
1	55.54	55.54	82.63	82.63
2	44.07	99.61	2.95	85.59
3	0.28	99.89	7.29	92.88

2.5. HS-SPME-GC-MS experimental procedure

10 mL of final standards or mineral water samples were put into 20 mL glass vials for SPME. A fibre coated with DVB/CAR/PDMS, 50/30 μm film thickness, was used for HS-SPME. Before the first use and before each analysis, each fibre was conditioned according to the specifications of the manufacturer at a temperature of 270 °C for 30 min.

The fibre was exposed to the headspace for a certain extraction time at a certain extraction temperature, after an incubation time of 2 min. The incubation and extraction steps were carried out with constant stirring. An extraction temperature of 80 °C and an extraction time of 45 min were ultimately chosen. Figures of merit were studied and target compounds were determined in natural and flavoured mineral water samples with these latter values for the HS-SPME extraction.

The desorption time was set at 5 min. The injection port temperature was 270 °C; the depth of penetration of the needle in the injector was 27 mm. Helium was used as the carrier gas at a flow rate of 1.3 mL min^{-1}

Table 1
Information regarding the bottled natural and flavoured mineral water samples, packaging, and the concentration found of each analyte (with 95 % confidence interval).

Sample code	Packaging					Water	Flavour	Found concentration (μg L^{-1})			
	Material	Volume (mL)	Colour	Recycled (%)	Cap colour			DEP	DBP	BBP	DEHP
W1	RPET	750	Blue	100	White	Still	Without	n.d.	n.d.	n.d.	n.d.
W2	PET	330	Uncoloured	–	Red	Still	Without	n.d.	n.d.	n.d.	n.d.
W3	Plastic[a]	1500	Uncoloured	–	Green	Still	Apple	n.d.	n.d.	n.d.	n.d.
W4	Plastic[a]	1500	Uncoloured	–	Yellow	Still	Lemon	n.d.	n.d.	n.d.	n.d.
W5	Plastic[a]	500	Green	–	Yellow	Sparkling	Lemon	n.d.	n.d.	n.d.	n.d.
W6	PET	2000	Uncoloured	–	Red	Still	Without	n.d.	n.d.	n.d.	n.d.
W7	RPET	1500	Blue	50	White	Still	Without	n.d.	n.d.	n.d.	n.d.
W8	RPET	500	Uncoloured	25	Grey	Still	Without	n.d.	n.d.	n.d.	n.d.
W9	Plastic[a]	1500	Uncoloured	–	Green	Still	Apple	n.d.	n.d.	n.d.	n.d.
W10	Plastic[a]	1500	Uncoloured	–	Yellow	Still	Lemon	n.d.	n.d.	n.d.	n.d.
W11	Plastic[a]	500	Uncoloured	–	Grey	Sparkling	Lemon	n.d.	n.d.	n.d.	n.d.
W12	Plastic[a]	1000	Green	–	Green	Sparkling	Lime	n.d.	n.d.	n.d.	n.d.
W13	RPET	500	Uncoloured	50	Blue	Still	Without	n.d.	n.d.	n.d.	n.d.
W14	Glass	250	Uncoloured	–	Grey metal	Sparkling	Without	n.d.	0.28 ± 0.04	n.d.	n.d.
W15	Plastic[a]	1250	Blue	–	Green	Still	Apple	n.d.	0.35 ± 0.03	n.d.	n.d.
W16	Plastic[a]	1250	Blue	–	Yellow	Still	Lemon	n.d.	0.28 ± 0.04	n.d.	n.d.
W17	Plastic[a]	500	Uncoloured	–	Grey	Sparkling	Sweet-orange and black raspberry	n.d.	0.27 ± 0.04	n.d.	n.d.
W18	Plastic[a]	1500	Uncoloured	–	Green	Still	Apple	n.d.	n.d.	n.d.	n.d.
W19	Plastic[a]	1250	Blue	–	Yellow	Still	Lemon	n.d.	n.d.	n.d.	n.d.
W20	Plastic[a]	500	Uncoloured	–	Grey	Sparkling	Tangerine and strawberry	n.d.	0.29 ± 0.09	n.d.	n.d.
W21	RPET	500	Uncoloured	100	Red	Still	Without	n.d.	n.d.	n.d.	n.d.
W22	Plastic[a]	500	Green	–	Green	Sparkling	Without	n.d.	n.d.	n.d.	n.d.

RPET: recycled polyethylene terephthalate; PET: polyethylene terephthalate.
[a] not indicated; n.d. not detected value.

and the initial pressure was set at 69.8 kPa.

The oven temperature was 40 °C for 5 min after injection and then was increased at a rate of 20 °C min^{-1} to 250 °C, which was held for 1 min. That temperature was ramped again at a rate of 10 °C min^{-1} to 280 °C, which was held for 1 min. Then a post-run step was carried out at 300 °C for 3 min. The total run time was 23.50 min (with the post-run step).

After each desorption process, the fibre was cleaned-up at 270 °C for 6 min.

The mass spectrometer operated in the electron impact (EI) ionization mode at 70 eV. The transfer line temperature was set at 300 °C, the ion source at 230 °C and the quadrupole at 150 °C. After a solvent delay of 13.1 min, the data were acquired in single ion monitoring mode, with 30 ms of ion dwell time, using five time acquisition windows: i) from 13.10 to 14.19 min for DEP (diagnostic ions: 105, 121, 132, 149, 177); ii) 14.20–15.09 min for IS (diagnostic ions: 80, 153, 171, 209, 227); iii) 15.10–15.49 min for DBP (diagnostic ions: 104, 121, 149, 205, 223); iv) 15.50–18.39 min for BBP (diagnostic ions: 91, 104, 149, 206, 238); and v) from 18.4 to 23.50 min for DEHP (diagnostic ions: 71, 149, 167, 207, 279).

At the beginning, middle, and end of each chromatographic sequence, system blanks (vials with no solution) were injected to control the clean-up of the whole HS-SPME-GC-MS system.

Plastic consumables were reduced as much as possible and the laboratory glassware used throughout the work was thoroughly cleaned twice with three solvents (n-hexane, acetone and methanol, in that order).

2.6. Software

MSD ChemStation version D.02.00.275 (Agilent Technologies, Inc.) with Data Analysis software was used for scan control and data acquisition. The TriPlus Sampler version 1.6.9 SPME (Thermo) was used to control the TriPlus autosampler. PLS_Toolbox software, version 9.0 Eigenvector Research Inc. (Wenatchee, WA, USA, 2021), for use with MATLAB, version 9.10.0.1739362 (R2021a, Mathworks, Inc., Natick, MA, USA, 2021), was used to perform the PARAFAC2 decompositions. STATGRAPHICS Centurion 19 (Statpoint Technologies, Inc., Herndon, VA, USA, 2019) was used to fit and validate regression models. DETARCHI program [34] was used to calculate the decision limit (CCα) and capability of detection (CCβ). The joint test on slope and intercept [35] and the model inversion to find the Pareto optimal front were computed with in-house programs written in MATLAB code.

3. Theory

3.1. PARAFAC and PARAFAC2 decompositions

The individual $I \times J$ data matrices, containing the abundances of the GC-MS analysis (I elution times and J m/z ratios), for each of K samples can be stacked to form a three-way $I \times J \times K$ array $\underline{\mathbf{X}}$, which can be decomposed with PARAFAC, a technique for multi-way data [32]. For F factors, the trilinear PARAFAC model is the one in Eq. (1):

$$x_{ijk} = \sum_{f=1}^{F} a_{if} b_{jf} c_{kf} + e_{ijk}, \quad i = 1,2,\ldots,I;\ j = 1,2,\ldots,J;\ k = 1,2,\ldots,K \tag{1}$$

When varying i, j, k, the columns vectors made with a_{if}, b_{jf} and c_{kf}, denoted $\mathbf{a}_f$, $\mathbf{b}_f$ and $\mathbf{c}_f$, contain the loadings on the chromatographic, spectral and sample profiles, respectively, of the f-th compound. In Eq. (1), e_{ijk} denotes the different residuals of the decomposition. It is said that GC-MS data are trilinear if the experimental three-way array is compatible with Eq. (1).

As for the validation with $F \geq 2$, the CORCONDIA (core consistency diagnostic [36]) index measures the trilinearity degree of an experimental three-way array. If the three-way array is trilinear, then a value of 100 for the CORCONDIA is obtained, which is its maximum. The PARAFAC decomposition is obtained by least squares, and it is a unique solution when the three-way array is trilinear and the appropriate number of factors has been chosen for the PARAFAC model [31].

PARAFAC2 is used to correct deviations from the assumed trilinearity (a typical situation in the case at hand is the occurrence of shifts in the retention time of the analytes from one sample to another [37,38]). The difference is that PARAFAC2 applies the same profiles ($\mathbf{b}_f$, f = 1, …,F) along the spectral mode, while allowing variations from one matrix to another in the chromatographic mode. This is formalised in the PARAFAC2 model in Eq. (2) with the addition of the superscript k to account for the dependence of the chromatographic profile on the k-th sample.

$$\mathbf{X} = (x_{ijk}) = \left(\sum_{f=1}^{F} a_{ij}^{k} b_{jf} c_{kf} + e_{ijk} \right), \quad i = 1,2,\ldots,I;\ j = 1,2,\ldots,J;\ k = 1,2,\ldots,K \tag{2}$$

One of the advantages of PARAFAC and PARAFAC2 models is the availability of the Q and T^2 (Q-residuals and Hotelling's T^2) statistics. By setting a probabilistic threshold value (e.g. at 95 % confidence level), outlier samples can be identified as those that exceed these thresholds. In that case, the samples are rejected and the PARAFAC or PARAFAC2 models are recalculated.

Another property is the unequivocal identification (and subsequent quantification with a calibration model) of the analyte of interest even in the presence of unknown interferents, because these interferent(s) will appear as new factor(s) in the PARAFAC or PARAFAC2 models due to the second-order advantage. A mathematical explanation on how PARAFAC2 decomposition works to ensure unequivocal identification can be found in the Annex to Ref. [26]. This property is particularly useful in a single-quadrupole GC-MS analysis, because of the possible presence of coeluents with the same m/z ratios as the compounds of interest.

3.2. PLS2 model

To relate CMP and CQAs, a PLS2 regression method has been used, which is widely used in chemometrics [39]. For the common situation in AQbD of having multiple responses (multivariate CQAs), PLS2 is used because it is a more flexible method [4,6,40] than OLS, which has to fit each individual single response in CQAs separately.

To formalize the PLS2 model, let $\mathbf{X}_{NV}$ denote the N x V data matrix whose rows are the setttings of the V CMPs (predictor variables) selected for each of the N experimental essays. On the other hand, $\mathbf{Y}_{NW}$ is an N x W matrix, where again the rows contain the W CQAs (response variables) that qualify the corresponding N experimental results.

Therefore, after the N experiments have been carried out, a training set $\{\mathbf{X}_{NV}, \mathbf{Y}_{NW}\}$ is available that gathers information about the functional relationship of the CQAs on the CMPs with respect to the analytical procedure to be modelled.

The PLS2 regression model is a compromise between variances and correlation, which is obtained by maximising a geometric mean or, equivalently, their product, Eq. (3).

$$max_{\mathbf{r},\mathbf{q}} \left\{ \mathrm{var}(\mathbf{Xr}) [\mathrm{corr}(\mathbf{Xr}, \mathbf{Yq})]^2 \mathrm{var}(\mathbf{Yq}) \right\} \text{ subject to } \|\mathbf{r}\| = \|\mathbf{q}\| = 1 \tag{3}$$

The maximisation of the product $\mathrm{var}(\mathbf{Xr})[\mathrm{corr}(\mathbf{Xr}, \mathbf{Yq})]^2\mathrm{var}(\mathbf{Yq})$ tends to look for directions with large variance in both **X**- and **Y**- blocks, avoiding those with small variance or little correlation with each other.

Regarding the selection of the number of latent variables, a cross-validation process is performed using the "leave one out" procedure. For each PLS2 model with A latent variables, the Root Mean Squared Error in Cross Validation (RMSECV) in Eq. (4) is calculated with the predicted values of the j-th response variable, $\mathbf{y}_j$, computed with the corresponding PLS2 model.

$$RMSECV_j(A) = \sqrt{\frac{\sum_{i=1}^{N}\left[y_{ij} - \hat{y}_{-ij}(A)\right]^2}{N}} \quad (4)$$

In Eq. (4), $\hat{y}_{-ij}(A)$ is the value estimated for the i-th coordinate of $\mathbf{y}_j$ when it is not in the training set.

The number of latent variables is selected as the value of A that reaches the minimum $RMSECV_j(A)$ for the **Y**-block.

The significance of the model was determined using a permutation test, which gives the probability that the PLS2 model is significantly different from one constructed with data taken at random from the training set {**X**, **Y**}. Specifically, the matrix of responses is repeatedly and randomly reordered, and a PLS2 model is fitted for each permutation. In this way, the distribution of rows in the **Y**-block is preserved, but they do not necessarily correspond to the linked row vector in the **X**-block. This procedure, thus, computes the probability that the original model is not significantly different from one generated by randomly shuffling the response **Y**.

4. Results and discussion

The AQbD methodology was applied to the extraction process of phthalates by HS-SPME-GC-MS. The CMPs were the extraction time and temperature. A standard solution with a concentration of 4, 0.7, 50, and 50 μg L^{-1} of DEP, DBP, BBP, and DEHP, respectively, and 0.4 μg L^{-1} of IS, was used. This solution was analysed at three different extraction temperatures (40, 60 and 80 °C) and six different extraction times (5, 20, 35, 50, 65 and 80 min). This allows for a comprehensive study of a large experimental region. Fig. 1 shows some of the chromatograms recorded in the study; it can be seen how the peaks of the target analytes in each chromatogram change with the extraction time and temperature used.

Each analysis was carried out in triplicate (i.e. a total of 54 analyses) over 9 days. One replicate of each temperature was measured each day, with a system blank at the beginning and end of each sequence, giving a total of 72 samples.

As mentioned above, the method of analysis was to be applied to the determination of phthalates in commercial mineral water samples, which may contain unexpected and unknown interferents. For this reason, three-way methods such as PARAFAC or PARAFAC2 were used, which have already proven to be very useful in such cases [30], and which also allow the unequivocal identification of the compounds, as required by the regulations [33]. In particular, PARAFAC2 was used due to the slight shifts in the chromatographic peaks that were observed [26, 41].

The GC-MS data from each chromatographic acquisition window, after baseline correction, were arranged in an $I \times J \times K$ data array $\underline{\mathbf{X}}$. The first way or mode is the chromatographic one (I elution times), the second one is the spectral way (J diagnostic ions) and the third is the sample way (K samples). The different data cubes obtained and their size are shown in Table S2 in the supplementary material.

These data arrays were decomposed using PARAFAC2, imposing the non-negativity constraint on the three profiles. The possible presence of outliers was checked using the Q and T^2 statistics (95 % confidence level), as in the rest of the paper. The characteristics of the PARAFAC2

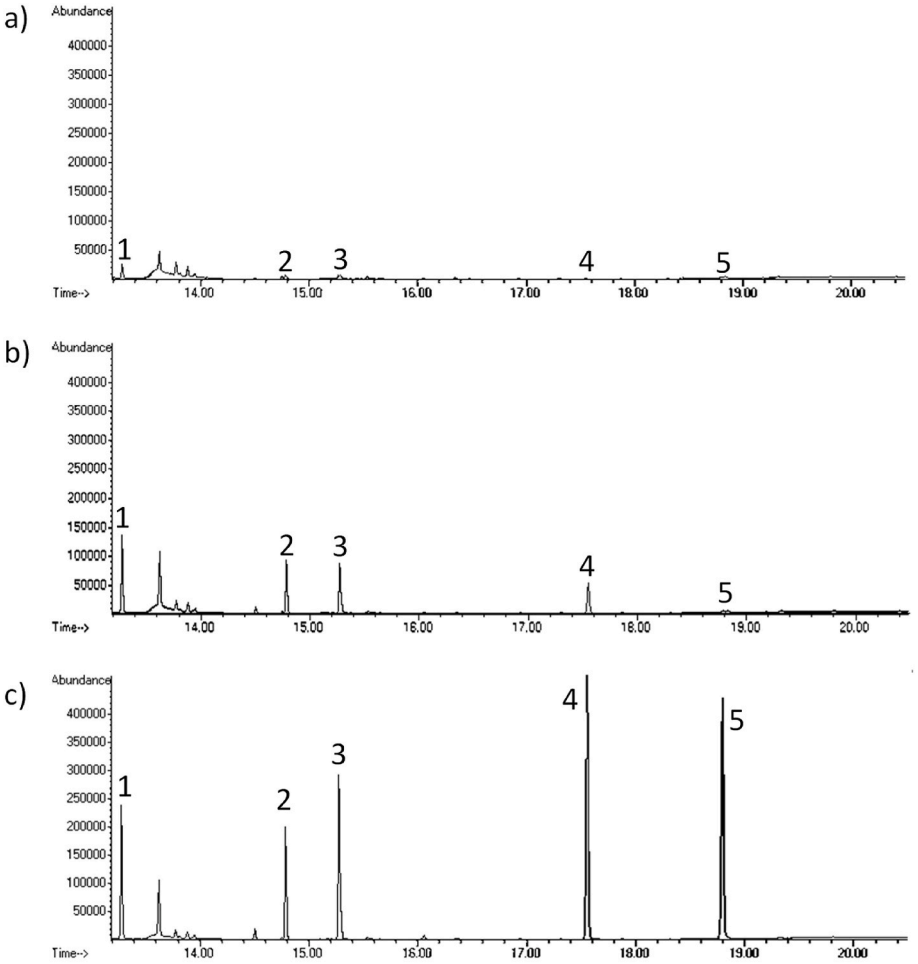

Fig. 1. Chromatograms obtained at different extraction temperature and extraction time: a) 40 °C and 80 min; b) 60 °C and 80 min; c) 80 °C and 50 min. Analytes: 1) DEP; 2) IS; 3) DBP; 4) BBP; and 5) DEHP.

models obtained are also shown in Table S2 in the supplementary material. In this part of the study, all the models have at least two factors, indicating the presence of a coeluent compound that shares *m/z* ratios with the target compound [42], even in the IS used. In all cases the CORCONDIA index is 100 % and the explained variance is greater than 99.98 %.

Fig. 2a–c shows the loadings obtained for DEP, the PARAFAC2 model explains the 100 % of the variance with 2 factors. The chromatographic profiles in Fig. 2a show how the chromatographic peaks of the DEP (in blue) and an interferent (in green) overlap and share *m/z* ratios, as can be seen in the spectral profile of each factor in Fig. 2b. The sample loadings in Fig. 2c show that the higher the extraction temperature and the longer the extraction time, the higher the loadings obtained for the DEP factor (in blue). The same effect is also observed for the loadings of the interferent (in green).

Once the PARAFAC models have been obtained, it is necessary to ensure the unequivocal identification of the target compounds through the chromatographic and spectral profiles of the corresponding factors. The requirements for the identification of analytes, taking into account the analysis of food contact materials [33], state that the relative retention time of the compound must correspond to that of the reference sample within a tolerance of ±0.5 %. For diagnostic ions, this regulation states that at least 3 *m/z* ratios should comply with the stated identification conditions.

In order to carry out the unequivocal identification, 9 samples are analysed: a system blank at the beginning and end of the sequence, and 7

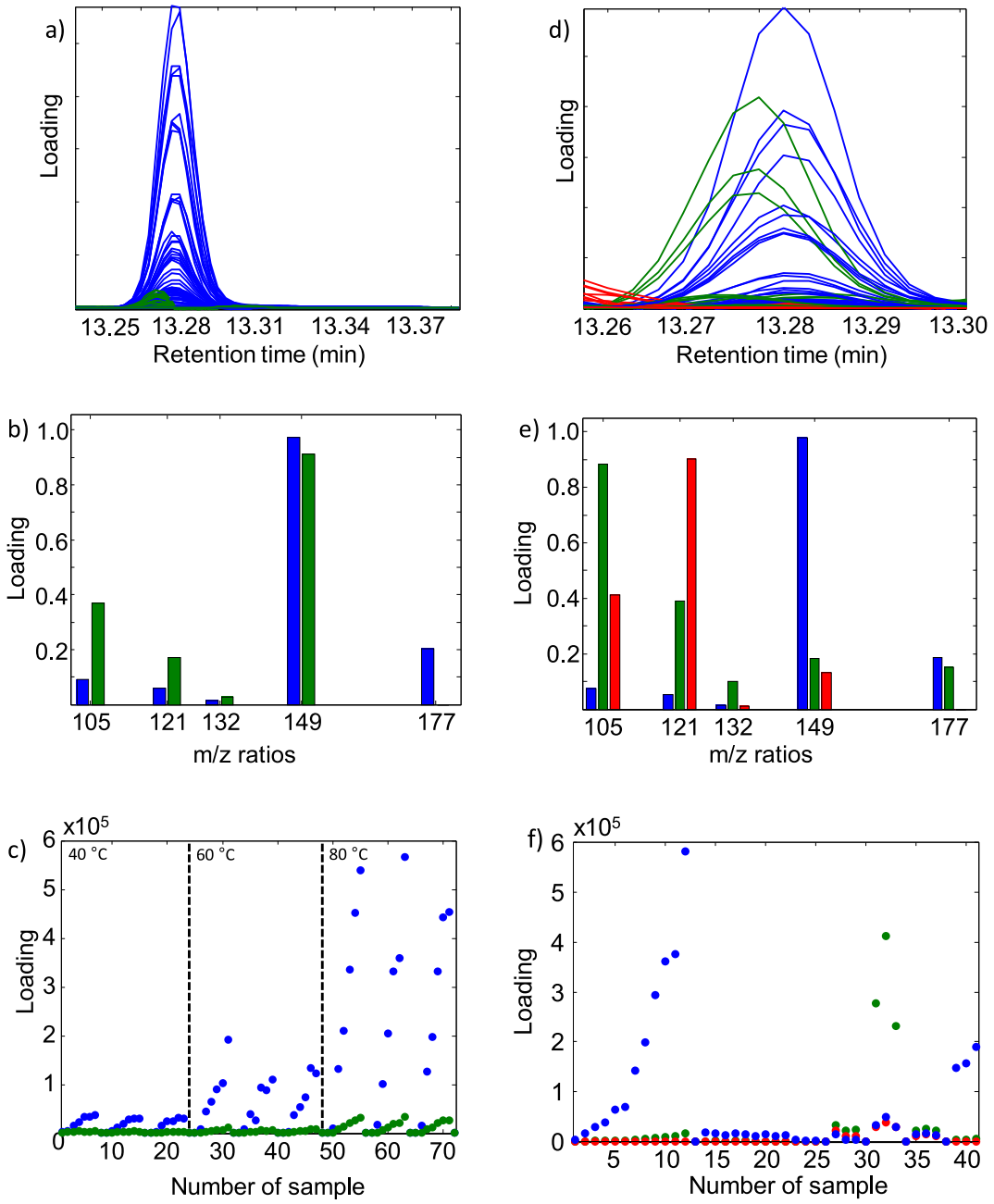

Fig. 2. Chromatographic (a, d), spectral (b, e) and sample (c, f) loadings of the PARAFAC2 models for DEP (DEP in blue and interferents in green and red). Profiles in a), b), c) correspond to the extraction temperature and time studied (replicate 1: samples 2–8, 26–31 and 50–55; replicate 2: samples 10–15, 34–39 and 58–63; replicate 3: samples 18–23, 42–47 and 66–71; the remaining samples are system blanks). Profiles d)-f) correspond to the bottled mineral water analysis (calibration set: samples 2–12; system blanks: 1, 13, 26, 30, 34, 38; samples 39-41 have different concentrations of IS and a fixed level of the target analytes; the remaining samples are water samples, from W6 to W12, in triplicate).

reference standards, 4 of them with different concentration levels of the target analytes (covering the wide range of concentrations used) and a fixed level of IS, and the other 3 with different concentrations of IS and a fixed level of the target analytes. These data are arranged in data cubes that are decomposed using PARAFAC2; the characteristics of the cubes and the PARAFAC2 models are shown in Table S2 in the supplementary material (in the rows about tolerance intervals). In this case, the models for DEP and BBP had 1 factor, 2 factors for DBP and DEHP, and 3 factors for the IS. The CORCONDIA index was always greater than 99 % and the explained variance was greater than 99.8 %.

The tolerance intervals obtained from the chromatographic and spectral profiles [33] are shown in Table S3 in the supplementary material. The table also shows the values for relative retention times and relative abundances for this study. It can be seen that for all compounds the requirements set out in the regulation [33] are fulfilled, so it can be concluded that the compounds are unequivocally identified.

4.1. Selection of the experimental conditions: fitting and inversion of a PLS2 model

To obtain the model relating CMPs and CQAs in the experimental space, a PLS2 was first constructed with autoscaled predictor variables in **X** (54 × 5) and five responses in **Y** (54 × 5). There are thus five responses corresponding to the five analytes (DEP, IS, DBP, BBP and DEHP) and five predictor variables used to construct the model: temperature (Ti), time (ti), and squared and cross-product terms, Ti × ti, Ti^2, ti^2. Therefore, the model postulated is the quadratic model in time and temperature as nonlinear behaviour is expected. The number of rows corresponds to the 54 experiments with the three extraction temperatures (40, 60, and 80 °C) and the six extraction times (5, 20, 35, 50, 65 and 80 min) for which the fibre was kept at each temperature. The specific CQAs used to characterise the compounds are the normalised sample loadings of the four phthalates and the internal standard (DEP, IS, DBP, BBP, and DEHP), obtained by PARAFAC2 decomposition, which are linearly related to the concentration.

The PLS2 model, with five latent variables, explains 90.7 % of the variance of all five responses. The contour levels of the surfaces fitted for each of the five analytes are depicted in Fig. 3a–e. In all of them, short times and the lowest temperature give expected loadings practically zero for all the analytes, therefore useless experimental conditions. Fig. 1a shows how the chromatographic peaks of the analytes are very low in abundance and practically not visible, even when extracted for 80 min. Consequently, to better explore the region of interest, another PLS2 model was fitted in a reduced experimental domain, made up with the new ranges of temperature and time, between 60 and 80 °C and between 35 and 80 min, respectively.

The new, more local, PLS2 model is fitted with the autoscaled **X** (24 × 5), and the corresponding responses in **Y** (24 × 5), which were centred. Three latent variables were chosen by cross-validation. In addition, to validate this choice, the p-value of the permutation test was less than 0.05, that is, the model with three latent variables is significant at the 95 % confidence level.

Table 2 shows the percentage of explained variance in the predictors **X,** and in the responses **Y**, when adding latent variables in the PLS2 model. Overall, the variability in **X** is explained and almost 93 % of that of **Y**, improving the percentage of the previous model, as expected for a more local model.

The five responses are more or less equally well fitted, Table 3 contains the individual coefficient of determination and coefficient of determination in crossvalidation, estimated with the leave one out procedure. The similarity between the explained variance in fitting and prediction, both high, indicates highly predictive models, an essential property for the inversion.

Fig. 3 shows that, for the diverse times and temperatures, the expected loadings vary differently depending on the response, although they all increase as the point (80, 80) is approached. In fact, it is not necessary to have maximum values of loadings in order to obtain experimental signals of sufficient quality to carry out the experiment, as the experiments already performed have shown: loadings around 0.20 are sufficient to obtain good chromatograms. This observed property can be used to look for alternative experimental conditions, other than (80, 80), that reduce the time of each analysis.

Considering DEHP, the analyte with the lowest sensitivity, the chromatogram in Fig. 1b (extraction temperature 60 °C and extraction time 80 min) shows a tiny peak (peak 5) that corresponds to a loading of 0.1. On the other hand, the chromatogram in Fig. 1c, obtained with an extraction temperature of 80 °C and an extraction time of 50 min, shows a much higher peak for DEHP, corresponding to a loading around 0.2.

Then, the PLS2 model should be inverted. The inversion is computational as there is no solution for the algebraic inversion due to the terms of degree 2 in the model [43]. Briefly, the method of inversion, explained in Ref. [40], consists of looking for the experimental conditions that give a compromise solution in terms of the desired or target values of the responses, while guaranteeing that these experimental conditions are inside both the experimental domain and the PLSbox that includes the part of the latent space and the part of the residual space bounded, in our case, by the 95 % confidence level for the Q and T^2 statistics.

In the present case, to better describe the individual behaviour of each compound, the target value defined for the inversion corresponds to half the observed range in the experiments conducted, that is, to look for loadings half the range of the experimental loadings in **Y**, that is, $\mathbf{y}_{target}$ = (0.20, 0.21, 0.22, 0.26, 0.25) for the responses in the order of Fig. 3, namely, DEP, IS, DBP, BBP, and DEHP.

Fig. 4 shows the parallel coordinates plot of the solutions found. In the graph, each 'solution' (consisting of experimental conditions and their expected responses) is represented by the sequence of its coordinate values plotted against the coordinate indices. The first two coordinates correspond to the experimental conditions (time and temperature), the last five to the expected values of the responses (the loadings), also in the order of Fig. 3. These are linked following the broken line connecting the corresponding individual solutions.

To avoid the different scales that would obscure the interpretation, all values are range-scaled and their corresponding minimum and maximum values are at the bottom and top of their respective coordinates. It can be seen how the same length in the vertical direction of Fig. 4 represents 32.88 min in the first coordinate, but only 6.83 °C in the second. The same applies to the loadings that, overall, vary between 0.20 and 0.27 for the first three responses and are slightly different for the last two phthalates.

The individual values targeted for each response are obtained in at least one of the solutions, but not together (otherwise, there would be a single line in the plot). This is an indication of the conflict among responses when trying to get close to $\mathbf{y}_{target}$, which is superimposed on the graph with a dashed green line, obviously only in the coordinates of the responses.

In fact, the solutions found are an estimate of the Pareto optimal front for the five responses (the five CQAs), that is, the set of solutions that are the best in at least one of the objectives, while maintaining the rest of them in their best value in the conditions stated. In other words, when moving along the front, no single CQA can be improved without worsening another CQA. In this sense, Fig. 4 shows the different behaviour of each individual response, in relation to the others. For example, the dotted blue line at the bottom of Fig. 4 corresponds to the solution with the target value for DEP (0.20), and also for IS (0.21), but it is seen how this is related to the lowest values for the loadings of BBP and DEHP, the most striking, graphically far from their target values (0.26 and 0.25), with the loading for DBP also less than desired (0.22). Following the broken line to the first coordinates, these loadings are expected with 45 min at around 78 °C (approximately a quarter of the line below 80 °C in the second coordinate).

On the other extreme, the line in cyan is the solution approaching the

target value of BBP and not very far from the one of DEHP (the last two compounds), obtained with 55 min at 80 °C. Similar loadings can be expected, dark blue line, by increasing the time up to 78 min and reducing the temperature until 73 °C, reflecting the opposite behaviour of the experimental conditions on the responses but quantifying this 'opposition' on the compromise achievable among responses.

With regard to this compromise, the red line (45 min, 80 °C), compared to the cyan line, shows that about 10 min can be saved per experiment without being very far from the target values (the first three loadings are close to their target values, the last two a little further, but not to the most extreme). Also, when comparing the 45 min with the theoretical maximum of 80 min, the time is almost halved, which represents a significant saving in time and energy.

4.2. Experimental verification of the selected conditions

To experimentally verify the control method parameters found in the inversion of the PLS2 prediction model, a standard solution with concentrations equal to those used to obtain these parameters was analysed in quintuplicate. Fig. 5a shows the chromatogram recorded for one of these replicates, where the analytical quality of all the peaks of the chromatogram obtained can be corroborated.

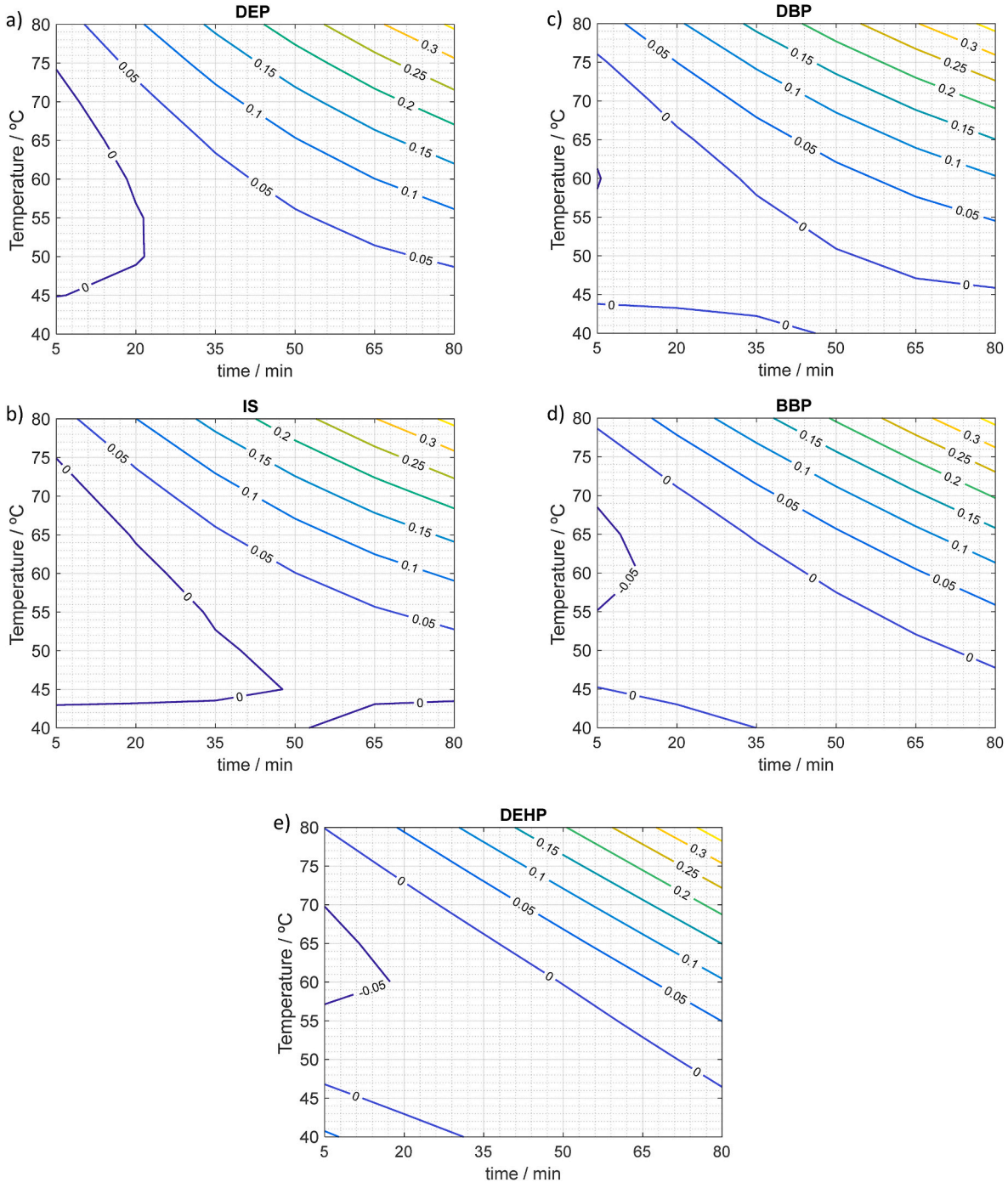

Fig. 3. Level curves of the surface fitted for the five analytes (a) DEP, b) IS, c) DBP, d) BBP, e) DEHP) in the domain of the CMPs (time and temperature of extraction used in the HS-SPME procedure).

Table 3
Coefficient of determination and coefficient of determination in cross-validation for the five responses fitted with PLS2. Y_1: DEP; Y_2: IS; Y_3: DBP; Y_4: BBP and Y_5: DEHP.

	Y_1	Y_2	Y_3	Y_4	Y_5
R^2 (%)	96.58	95.42	94.98	88.70	91.75
R^2_{CV} (%)	94.68	93.29	92.88	83.65	87.93

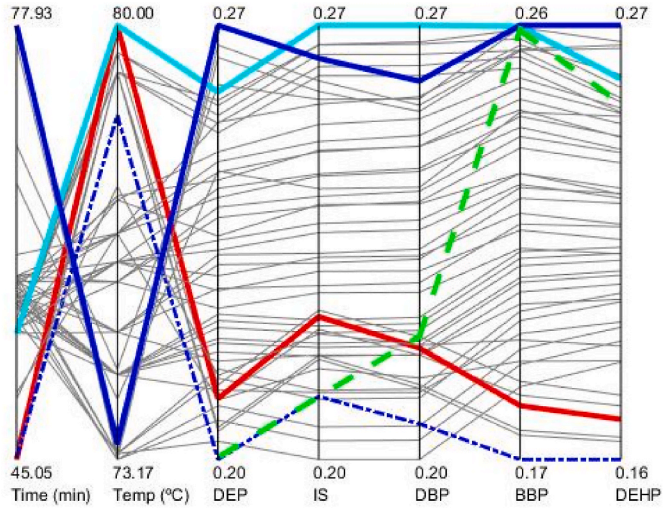

Fig. 4. Parallel coordinates plot of the Pareto optimal front found when inverting the PLS2 model to approach the dashed green line. See text for explanation about the rest of the coloured lines.

A successful PARAFAC2 decomposition requires different concentrations of all the compounds under consideration. Therefore, these analyses were carried out together with those of 14 other standards. 11 of these standards are used below, in Section 4.3, for the determination of some figures of merit, while in the remaining 3, only the concentration of the IS was varied in order to help the PARAFAC2 decomposition to model the IS contribution. In addition, 4 system blanks were analysed, giving a total of 23 analyses.

The GC-MS data from the 23 analyses were arranged in the arrays with the dimension shown in Table S2 in the supplementary material (rows corresponding to verification and figures of meirt), and the corresponding PARAFAC2 models were obtained, models with the same characteristics as those obtained for the estimation of the tolerance intervals. In all cases, both the relative retention time and the relative abundance for the diagnostic ions (Table S3 in the supplementary material) were within the corresponding tolerance intervals and therefore all compounds were unequivocally identified according to the regulation [33].

Table 4 contains the values of the expected loadings for each compound at the selected extraction time and temperature (45 °C, 80 min, in red in Fig. 4), as well as the values obtained experimentally from the five replicates analysed. As can be seen, the expected and experimental values are significantly equal, showing the success of the PLS2 model inversion and ensuring that the analytical method meets the requirements set to obtain an acceptable chromatogram.

4.3. Figures of merit of the analytical procedure

From the loadings obtained in Section 4.2 for the 11 standard solutions, some figures of merit of the analytical method were estimated for each target analyte: linear range, accuracy (trueness and intermediate precision), CCα and CCβ. As usual in chromatography for quantitative analysis, standardised signals were used. For this purpose, standardised loadings were calculated by dividing each loading by that of the internal standard [44].

The linear dynamic ranges found for each target compound are shown in Table 5. Outlier detection prior to fitting a least squares model was always performed by robust least squares (least median squares, LMS) regression [45].

Trueness and intermediate precision were determined from the accuracy lines, the parameters of which are shown in Table 5. Trueness was assessed from the joint hypothesis test (null hypothesis: slope = 1 and intercept = 0), whose *P*-values are greater than 0.05. Therefore, at 5 % significance level, there is not evidence to reject the null hypothesis and, therefore, it is concluded that the trueness is fulfilled. The residual standard deviation of the accuracy lines (Table 5) is considered as an estimate of the intermediate precision in the concentration range analysed.

The values found for the decision limit and the detection capability are also shown in Table 5. The former vary from 0.13 μg L^{-1} for DBP to 3.73 μg L^{-1} for DEHP for a false positive probability of 0.05, and the latter from 0.26 μg L^{-1} for DBP to 7.20 μg L^{-1} for DEHP when the probabilities of false positive and false negative are both fixed at 0.05.

4.4. Analysis of bottled natural and flavoured mineral water samples

The 22 bottled mineral water samples in Table 1 were analysed using the extraction temperature and extraction time found in Section 4.1. The analyses were carried out over a period of 4 days. Each day, a set of 11 calibration standards was measured, followed by some bottled mineral water samples in triplicate (with system blanks at the beginning of the sequence and after each set of replicates of flavoured samples) and finally 3 samples with different IS concentrations. Fig. 5b–c shows the chromatograms obtained for a natural mineral water sample (W2) and a flavoured mineral water sample (W20). As can be seen in the figure, on the same scale, sample W20 has many more chromatographic peaks corresponding to the volatile compounds extracted by HS-SPME-GC-MS. Sample W2 has far fewer compounds retained in the SPME fibre.

The data arrays obtained from these samples, as well as the characteristics of the corresponding PARAFAC2 models built, are shown in Table S2 in the supplementary material. The PARAFAC2 models have CORCONDIA values between 93 and 100 %, and all of them explain more than 99.8 % of the variance, except for one of the models built for the IS, where the CORCONDIA is 72 %, probably due to the appearance of an additional interferent (this model had 4 factors, while the others for the IS had only 3). In all cases, all compounds were unequivocally identified, as can be seen in Table S3 in the supplementary material; all relative retention times are within the relevant tolerance ranges, as are the relative abundances of at least 3 diagnostic ions for each compound, as regulation establishes [33].

Fig. 2d–f shows the loadings of the PARAFAC2 model obtained on day 2 for DEP, which explains 99.96 % of the variance with 3 factors. The factor related to DEP is shown in blue, and the other two correspond to interferents. The interferent in green, whose chromatographic peak completely overlaps with that of DEP (Fig. 2d), is found only in the three replicates of sample W11 (samples 31–33 in Fig. 2f). This shows the ability of PARAFAC2 to find this type of coeluent. In these determinations, the use of PARAFAC2 has been decisive, as it has made it possible to detect the presence of interferents and to differentiate their signal from that of the analytes of interest, without the need to vary the chromatographic method, thus making it possible to identify them unequivocally.

The results of the determinations are shown in the last four columns of Table 1. None of the target compounds were detected in most of the samples. Residues of only one of the four phthalates under study, DBP, were found in 5 of the 22 bottled mineral waters analysed, 4 of them flavoured mineral waters and one natural mineral water, with concentrations around 0.3 μg L^{-1}.

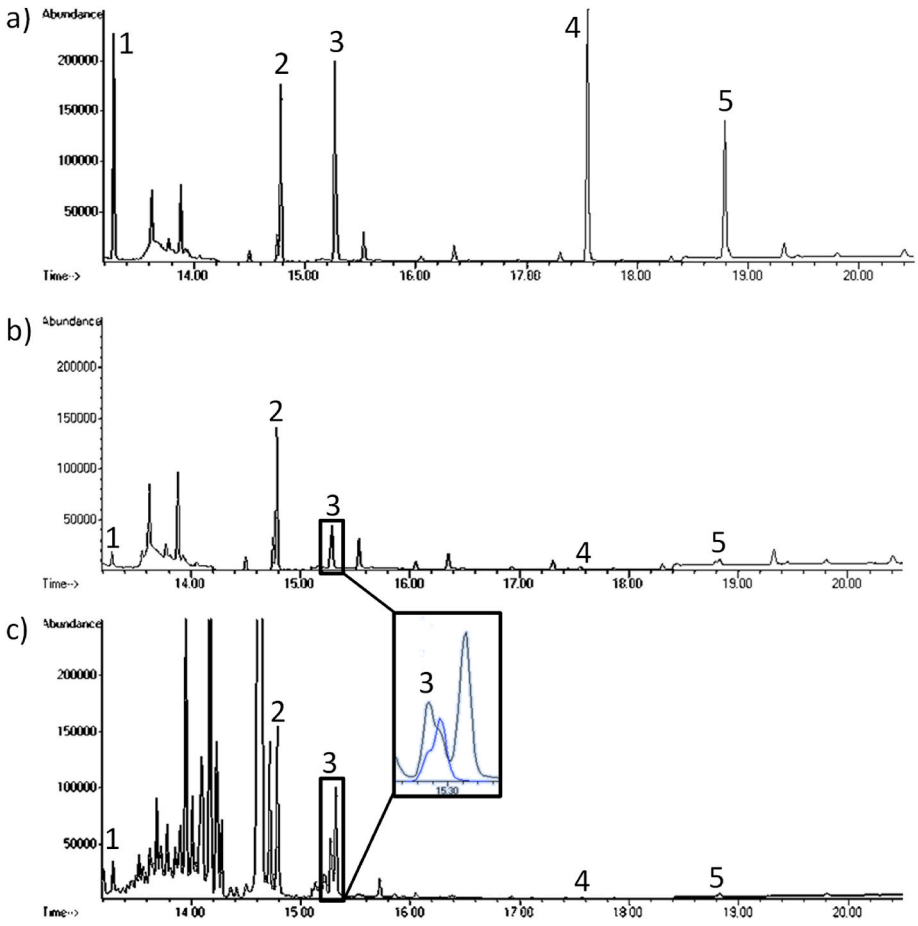

Fig. 5. Chromatograms obtained at 80 °C and 45 min: a) standard sample of the verification step; b) water sample W2; c) water sample W20. The black box shows a zoom corresponding to the DBP of images b and c (W2 in blue and W20 in black). Analytes: 1) DEP; 2) IS; 3) DBP; 4) BBP; and 5) DEHP.

As an example, the zoom in Fig. 5b–c shows the DBP peak (peak number 3) of sample W20 (in black), corresponding to a concentration of 0.29 μg L^{-1}, while for sample W2 (in blue) the presence of this compound is not detected, as PARAFAC2 has extracted the coeluent, hence the importance of its use to detect interferents.

Table 4
Expected and experimental (with their 95 % confidence intervals) loadings obtained for each compound at the selected extraction time and temperature.

Analyte	Expected loading	Experimental loading
DEP	0.209	0.185 ± 0.042
IS	0.223	0.186 ± 0.077
DBP	0.218	0.169 ± 0.070
BBP	0.184	0.200 ± 0.102
DEHP	0.172	0.196 ± 0.091

5. Conclusions

This work represents a novel application of AQbD to design a HS-SPME-GC-MS procedure. The procedure starts with a systematic coverage of the experimental domain, i. e. the region that contains viable values for the CMPs. The novelty is that the intended responses (CQAs) are defined in terms of sample loadings instead of, for example, peak area or resolution, which is the usual practice. In this way, it is possible to apply PARAFAC/PARAFAC2 methods that ensure the unequivocal identification of the analytes and allow the separation and extraction of the coeluents of the analyte of interest.

The inversion of a PLS2 model to obtain the intended quality chromatograms is the key to selecting the chosen solution. With the inversion, the resulting optimal experimental conditions give expected responses fulfilling the requirements with 45 min at 80 °C, almost

Table 5
Figures of merit of the analytical experimental procedure.

Analyte	Linear range (μg L^{-1})	Accuracy line					CCα (μg L^{-1})	CCβ (μg L^{-1})
		Intercept	Slope	r	s_{yx} (μg L^{-1})	*P*-value		
DEP	0–8	0.000	1.000	0.999	0.144	1.0000	0.62	1.20
DBP	0–2	0.001	0.999	0.994	0.078	0.9997	0.13	0.26
BBP	0–100	0.001	1.000	0.994	4.051	1.0000	1.82	3.52
DEHP	0–100	−0.002	1.000	0.997	3.402	0.9997	3.73	7.20

r: correlation coefficient; s_{yx}: standard deviation of regression; P-value when jointly testing {intercept = 0 and slope = 1}; CCα: decision limit; CCβ: capability of detection (α = β = 0.05).

halving the time of analysis.

The method developed was also validated and evaluated on real samples corresponding to bottled mineral waters, with very low concentrations of phthalates. Only DBP was found at concentrations around 0.3 μg L^{-1} in 5 out of the 22 samples analysed.

Although the method is used in a specific matrix, the proposed protocol for the application of AQbD with HS-SPME-GC-MS is completely general and would be applicable to complex matrices, provided that the experimental conditions are suitable for the determination.

CRediT authorship contribution statement

Lucía Valverde-Som: Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft, Writing – review & editing. **Ana Herrero:** Conceptualization, Methodology, Formal analysis, Writing – original draft, Writing – review & editing. **Celia Reguera:** Conceptualization, Methodology, Investigation, Writing – review & editing. **Luis Antonio Sarabia:** Conceptualization, Methodology, Software, Formal analysis, Writing – original draft, Writing – review & editing, Supervision. **María Cruz Ortiz:** Conceptualization, Methodology, Writing – review & editing, Supervision, Funding acquisition. **María Sagrario Sánchez:** Conceptualization, Methodology, Software, Formal analysis, Writing – original draft, Writing – review & editing, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that has been used is confidential.

Acknowledgments

The authors thank the financial support provided by the Consejería de Educación de la Junta de Castilla y León (JCyL) through project BU052P20, co-financed with European FEDER funds. Lucía Valverde-Som thanks JCyL for her postdoctoral contract through project BU052P20.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.talanta.2023.125265.

References

[1] P. Ramalingam, B. Jahnavi, QbD considerations for analytical development, in: S. Beg, M.S. Hasnain (Eds.), Pharmaceutical Quality by Design, Academic Press London Wall, London EC2Y 5AS, 2019, pp. 77–108, https://doi.org/10.1016/B978-0-12-815799-2.00005-8.

[2] S. Beg, J. Haneef, M. Rahman, R. Peraman, M. Taleuzzaman, W.H. Almalki, Introduction to analytical quality by design, in: S. Beg, M.S. Hasnain, M. Rahman, W.H. Almalki (Eds.), Handbook of Analytical Quality by Design, Academic Press, 2021, pp. 1–14, https://doi.org/10.1016/B978-0-12-820332-3.00009-1.

[3] S. Beg, M. Rahman, Analytical quality by design for liquid chromatographic method development, in: S. Beg, M.S. Hasnain, M. Rahman, W.H. Almalki (Eds.), Handbook of Analytical Quality by Design, Academic Press, 2021, pp. 87–97, https://doi.org/10.1016/B978-0-12-820332-3.00010-8.

[4] M.M. Arce, S. Ruiz, S. Sanllorente, M.C. Ortiz, L.A. Sarabia, M.S. Sánchez, A new approach based on inversion of a partial least squares model searching for a preset analytical target profile. Application to the determination of five bisphenols by liquid chromatography with diode array detector, Anal. Chim. Acta 1149 (2021), 338217, https://doi.org/10.1016/j.aca.2021.338217.

[5] M.C. Ortiz, L.A. Sarabia, A. Herrero, C. Reguera, S. Sanllorente, M.M. Arce, O. Valencia, S. Ruiz, M.S. Sánchez, Partial least squares model inversion in the chromatographic determination of triazines in water, Microchem. J. 164 (2021), 105971, https://doi.org/10.1016/j.microc.2021.105971.

[6] M.M. Arce, S. Sanllorente, S. Ruiz, M.S. Sánchez, L.A. Sarabia, M.C. Ortiz, Method operable design region obtained with a partial least squares model inversion in the determination of ten polycyclic aromatic hydrocarbons by liquid chromatography with fluorescence detection, J. Chromatog. A 1657 (2021), 462577, https://doi.org/10.1016/j.chroma.2021.462577.

[7] M.C. Ortiz, L.A. Sarabia, M.S. Sánchez, The inversion of multiresponse partial least squares models, a useful tool to improve analytical methods in the framework of analytical quality by design, Anal. Chim. Acta 1276 (2023), 341620, https://doi.org/10.1016/j.aca.2023.341620.

[8] L. Chiarentin, C. Gonçalves, C. Augusto, M. Miranda, C. Cardoso, C. Vitorino, Drilling into "quality by design" approach for analytical methods, Crit. Rev. Anal. Chem. (2023), https://doi.org/10.1080/10408347.2023.2253321.

[9] S. Orlandini, G. Hancu, Z.-I. Szabó, A. Modroiu, L.-A. Papp, R. Gotti, S. Furlanetto, New trends in the quality control of enantiomeric drugs: quality by design-compliant development of chiral capillary electrophoresis methods, Molecules 27 (2022) 7058, https://doi.org/10.3390/molecules27207058.

[10] R.S. Hanafi, M. Lämmerhofer, Quality-by-design approach for development of aqueous headspace microextraction GC-MS method for targeted metabolomics of small aldehydes in plasma of cardiovascular patients, Anal. Chim. Acta 1221 (2022), 340176, https://doi.org/10.1016/j.aca.2022.340176.

[11] K. Lee, W. Yoo, J.H. Jeong, Analytical method development for 19 alkyl halides as potential genotoxic impurities by analytical quality by design, Molecules 27 (2022) 4437, https://doi.org/10.3390/molecules27144437.

[12] K. Lee, W. Yoo, J.H. Jeong, Determination of three alkyl camphorsulfonates as potential genotoxic impurities using GC-FID and GC-MS by analytical QbD, Separations 9 (2022) 246, https://doi.org/10.3390/separations9090246.

[13] J. Freitas, P. Silva, P. Vaz-Pires, J.S. Câmara, A systematic AQbD approach for optimization of the most influential experimental parameters on analysis of fish spoilage-related volatile amine, Foods 9 (2020) 1321, https://doi.org/10.3390/foods9091321.

[14] J.N. Hahladakis, C.A. Velis, R. Weber, E. Iacovidou, P. Purnell, An overview of chemical additives present in plastics: migration, release, fate and environmental impact during their use, disposal and recycling, J. Hazard Mater. 344 (2018) 179–199, https://doi.org/10.1016/j.jhazmat.2017.10.014.

[15] L.E. Gray, J. Ostby, J. Furr, M. Price, D.N.R. Veeramachaneni, L. Parks, Perinatal exposure to phthalates DEHP, BBP, and DiNP, but not DEP, DMP, or DOTP, alters sexual differentiation of the male rat, Toxicol. Sci. 58 (2000) 350–365, https://doi.org/10.1093/toxsci/58.2.350.

[16] R. Akhbarizadeh, S. Dobaradaran, T.C. Schmidt, I. Nabipour, J. Spitz, Worldwide bottled water occurrence of emerging contaminants: a review of the recent scientific literature, J. Hazard Mater. 392 (2020), 122271, https://doi.org/10.1016/j.jhazmat.2020.122271.

[17] Q. Luo, Z.H. Liu, H. Yin, Z. Dang, P.X. Wu, N.W. Zhu, Z. Lin, Y. Liu, Migration and potential risk of trace phthalates in bottled water: a global situation, Water Res. 147 (2018) 362–372, https://doi.org/10.1016/j.watres.2018.10.002.

[18] L. Huang, X. Zhu, S. Zhou, Z. Cheng, K. Shi, C. Zhang, H. Shao, Phthalic acid esters: natural sources and biological activities, Toxins 13 (2021) 495, https://doi.org/10.3390/toxins13070495.

[19] W. Meng, H. Sun, G. Su, Plastic packaging-associated chemicals and their hazards - an overview of reviews, Chemosphere 331 (2023), 138795, https://doi.org/10.1016/j.chemosphere.2023.138795.

[20] M. Plotan, C. Frizzell, V. Robinson, C.T. Elliott, L. Connolly, Endocrine disruptor activity in bottled mineral and flavoured water, Food Chem. 136 (2013) 1590–1596, https://doi.org/10.1016/j.foodchem.2012.01.115.

[21] REGULATION (EC) No 1907/2006 of the EUROPEAN PARLIAMENT and of the COUNCIL of 18 December 2006 Concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), Establishing a European Chemicals Agency, Amending Directive 1999/45/EC and Repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as Well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC, OJ L 396 30.12, 2006.

[22] Commission Regulation (EU) No 10/2011 of 14 January 2011 on Plastic Materials and Articles Intended to Come into Contact with Food. OJ L12, 15.1, 2011.

[23] A. P S, V. Vinod, P.B. Harathi, A critical review on extraction and analytical methods of phthalates in water and beverages, J. Chromatogr. A 1675 (2022), 463175, https://doi.org/10.1016/j.chroma.2022.463175.

[24] J. González-Sálamo, B. Socas-Rodríguez, J. Hernández-Borges, Analytical methods for the determination of phthalates in food, Curr. Opin. Food Sci. 22 (2018) 122–136, https://doi.org/10.1016/j.cofs.2018.03.002.

[25] S. Garrigues, M. de la Guardia, Challenges in Green Analytical Chemistry, vol. 66, Royal Society of Chemistry, 2020.

[26] L. Valverde-Som, A. Herrero, C. Reguera, L.A. Sarabia, M.C. Ortiz, A new multi-factor multi-objective strategy based on a factorial presence-absence design to determine polymer additive residues by means of head space-solid phase microextraction-gas chromatography-mass spectrometry, Talanta 253 (2023), 124021, https://doi.org/10.1016/j.talanta.2022.124021.

[27] Y.L. Feng, J. Zhu, E. Sensenstein, Development of a headspace solid-phase microextraction method combined with gas chromatography mass spectrometry for the determination of phthalate esters in cow milk, Anal. Chim. Acta 538 (2005) 41–48, https://doi.org/10.1016/j.aca.2005.02.020.

[28] J. Maalouly, N. Hayeck, A. Kassouf, D.N. Rutledge, V. Ducruet, Chemometric tools to highlight possible migration of compounds from packaging to sunflower oils, J. Agr. Food Chem. 61 (2013) 10565–10573, https://doi.org/10.1021/jf402888e.

[29] J.D. Carrillo, C. Salazar, C. Moreta, M.T. Tenea, Determination of phthalates in wine by headspace solid-phase microextraction followed by gas chromatography-mass spectrometry: fibre comparison and selection, J. Chromatogr. A 1164 (2007) 248–261, https://doi.org/10.1016/j.chroma.2007.06.059.
[30] M.C. Ortiz, S. Sanllorente, A. Herrero, C. Reguera, L. Rubio, M.L. Oca, L. Valverde-Som, M.M. Arce, M.S. Sánchez, L.A. Sarabia, Three-way PARAFAC decomposition of chromatographic data for the unequivocal identification and quantification of compounds in a regulatory framework, Chemometr. Intell. Lab. 200 (2020), 104003, https://doi.org/10.1016/j.chemolab.2020.104003.
[31] M.C. Ortiz, L.A. Sarabia, M.S. Sánchez, A. Herrero, S. Sanllorente, C. Reguera, Usefulness of PARAFAC for the quantification, identification, and description of analytical data, in: A. Muñoz de la Peña, H.C. Goicoechea, G.M. Escandar, A. C. Olivieri (Eds.), Data Handling in Science and Technology, Elsevier, Amsterdam, 2015, pp. 37–81, https://doi.org/10.1016/B978-0-444-63527-3.00002-3.
[32] R. Bro, PARAFAC. Tutorial and applications, Chemometr. Intell. Lab. 38 (1997) 149–171, https://doi.org/10.1016/S0169-7439(97)00032-4.
[33] S. Bratinova, B. Raffael, C. Simoneau, Guidelines for Performance Criteria and Validation Procedures of Analytical Methods Used in Controls of Food Contact Materials, first ed., EUR 24105EN, Luxembourg, 2009. https://publications.jrc.ec.europa.eu/repository/handle/JRC53034. accessed July 2023.
[34] L.A. Sarabia, M.C. Ortiz, DETARCHI: a program for detection limits with specified assurance probabilities and characteristic curves of detection, TrAC-Trends Anal. Chem. 13 (1994) 1–6, https://doi.org/10.1016/0165-9936(94)85052-6.
[35] M.C. Ortiz, M.S. Sánchez, L.A. Sarabia, Quality of analytical measurements: univariate regression, in: S. Brown, R. Tauler, B. Walczak (Eds.), Comprehensive Chemometrics, Elsevier, Amsterdam, 2020, pp. 71–105, https://doi.org/10.1016/B978-0-12-409547-2.14869-3.
[36] R. Bro, H.A.L. Kiers, A new efficient method for determining the number of components in PARAFAC models, J. Chemometr. 17 (2003) 274–286, https://doi.org/10.1002/cem.801.
[37] H.A.L. Kiers, J.M.F. ten Berge, R. Bro, PARAFAC2 – Part I. A direct fitting algorithm for the PARAFAC2 model, J. Chemometr. 13 (1999) 275–294, https://doi.org/10.1002/(SICI)1099-128X(199905/08)13:3/4<275::AID-CEM543>3.0.CO;2-B.
[38] R. Bro, C.A. Andersson, H.A.L. Kiers, PARAFAC2 – Part II. Modelling chromatographic data with retention time shift, J. Chemometr. 13 (1999) 295–309, https://doi.org/10.1002/(SICI)1099-128X(199905/08)13:3/4<295::AID-CEM547>3.0.CO;2-Y.
[39] S. Wold, Chemometrics and bruce: some fond memories, in: B.K. Lavine, S. D. Brown, K.S. Booksh (Eds.), 40 Years of Chemometrics – from Bruce Kowalski to the Future, ACS Symposium Series, American Chemical Society, Washington, DC, 2015, pp. 1–13, https://doi.org/10.1021/bk-2015-1199.ch001.
[40] S. Ruiz, M.C. Ortiz, L.A. Sarabia, M.S. Sánchez, A computational approach to partial least squares model inversion in the framework of the process analytical technology and quality by design initiatives, Chemometr. Intell. Lab. 182 (2018) 70–78, https://doi.org/10.1016/j.chemolab.2018.08.014.
[41] A. Herrero, C. Reguera, M.C. Ortiz, L.A. Sarabia, Determination of dichlobenil and its major metabolite (BAM) in onions by PTV–GC–MS using PARAFAC2 and experimental design methodology, Chemometr. Intell. Lab. 133 (2014) 92–108, https://doi.org/10.1016/j.chemolab.2013.12.001.
[42] M.L. Oca, L. Rubio, L.A. Sarabia, M.C. Ortiz, Dealing with the ubiquity of phthalates in the laboratory when determining plasticizers by gas chromatography/mass spectrometry and PARAFAC, J. Chromatogr. A 1464 (2016) 124–140, https://doi.org/10.1016/j.chroma.2016.07.074.
[43] S. Ruiz, L.A. Sarabia, M.C. Ortiz, M.S. Sánchez, Residual spaces in latent variables model inversion and their impact in the design space for given quality characteristics, Chemometr. Intell. Lab. 203 (2022), 104040, https://doi.org/10.1016/j.chemolab.2020.104040.
[44] M.C. Ortiz, L.A. Sarabia, Quantitative determination in chromatographic analysis based on n-way calibration strategies, J. Chromatogr. A 1158 (2007) 94–110, https://doi.org/10.1016/j.chroma.2007.04.047.
[45] M.C. Ortiz, L.A. Sarabia, A. Herrero, Robust regression techniques: a useful alternative for the detection of outlier data in chemical analysis, Talanta 70 (2006) 499–512, https://doi.org/10.1016/j.talanta.2005.12.058.